Brandenburg University of Technology Cottbus - Senftenberg

8. October 2017

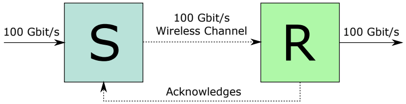
100 GBIT/S END-TO-END COMMUNICATION:

FLEXIBLE PROTOCOL PROCESSING ON MANYCORE NICS

Steffen Büchner, Jörg Nolte,
Lukasz Lopacinski, Rolf Kraemer

Project End2End100

CHALLENGES

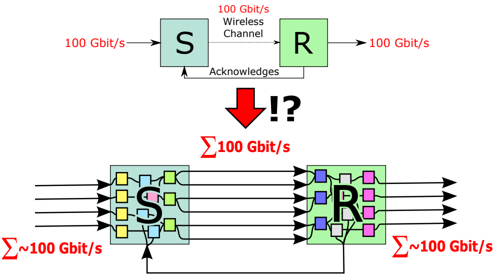


Desired throughput 100 Gbit/s and data-packet size of 1500 Byte:

- ⇒ We have to handle one data-packet every 120 ns.
- ⇒ The main memory access latency for a 64 Byte cache-line is ~96.6 ns[1].
- ⇒ The protocol processing has to be parallelized and offloaded into a manycore NIC!

[1] D. Molka, D. Hackenberg, R. Schöne, and W. E. Nagel, "Cache coherence protocol and memory performance of the intel haswell-ep architecture," in Parallel Processing (ICPP), 2015 44th International Conference on, Sept 2015, pp. 739–748.

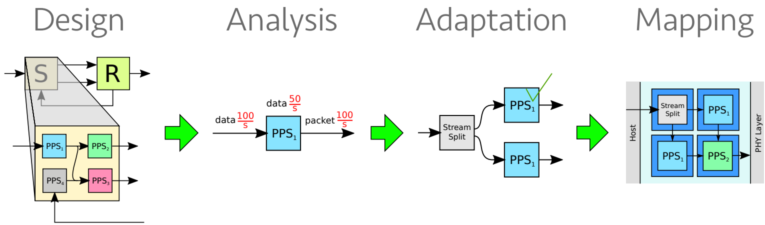
MANYCORE NIC



To be able to utilize a manycore processor for the protocol processing ...

- . . . we need easy to parallelize protocols.
- . . . we need a way to manage the parallel protocol processing.
- . . . we have to adapt the protocol processing to the available processing power.

A SOFT REAL-TIME STREAM PROCESSING DESIGN PROCESS

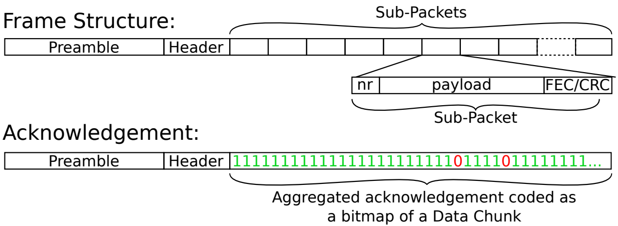


One design process . . .

- . . . without paradigm changes,
- . . . that allows us to adapt the protocol for different scenarios,
- . . . scale the protocol for different data-rates,
- . . . and map it on the manycore NIC.

DESIGN

AN EASY TO PARALLELIZE PROTOCOL STRUCTURE

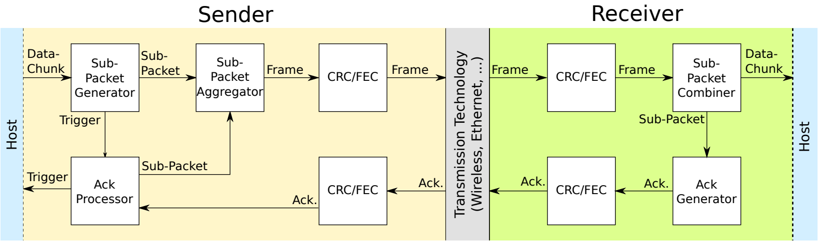


We have to combine . . .

- . . . large frames with small packets,
- . . . to minimize the protocol overhead as well as possible packet loss.

DESIGN

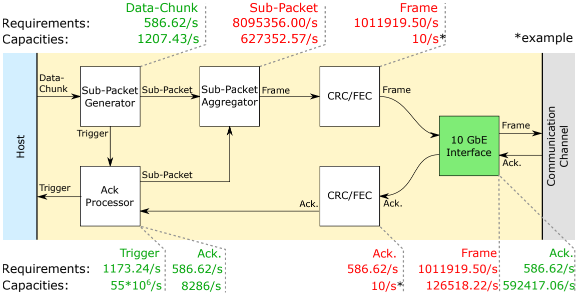
THE PROTOCOL PROCESSING ENGINE



A communication protocol is represented as a . . .

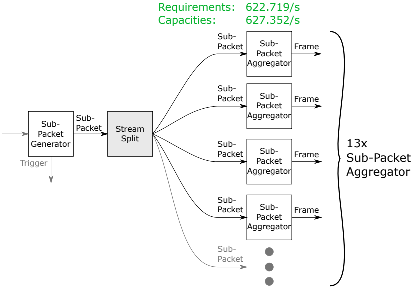
- . . . Protocol Processing Engine, which is composed of
- . . . Protocol Processing Stages, that fulfill the individual processing tasks.

ANALYSIS



- Each Protocol Processing Stage has soft real-time requirements that state how often this Protocol Processing Stage has to be executed.
- The hardware capacities state how often a Protocol Processing Stage can be executed on a single CPU.

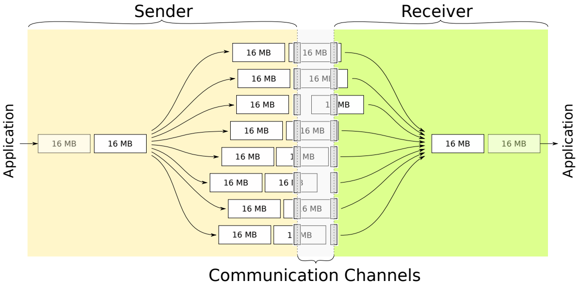
Brandenburg University of Technology Cottbus - Senftenberg



Stream-Operators are used to . . .

- . . . adapt the soft real-time requirements to the hardware's capacities
- . . . and to alter the protocol's behavior!

ADAPTATION
FOR STABLE THROUGHPUT



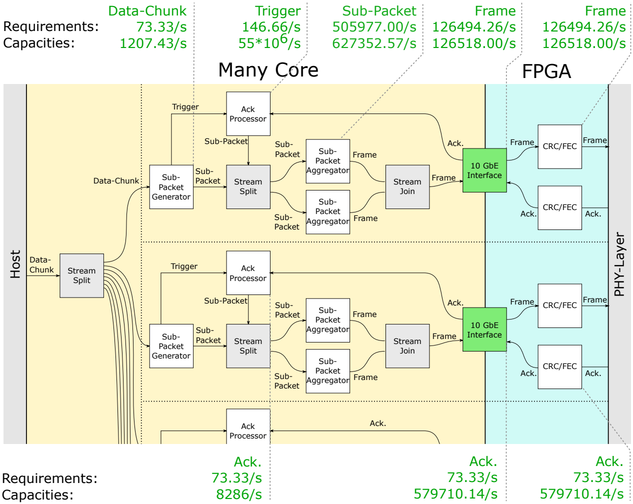
Several Data-Chunks are transmitted in parallel:

- Robust against packet loss
- Protocol overhead and retransmissions hidden by parallel pipelines
- Throughput expected to increase accordingly to the number of channels
- Latency is expected to stay constant

ADAPTATION

PROTOCOL PROCESSING ENGINE FOR STABLE THROUGHPUT

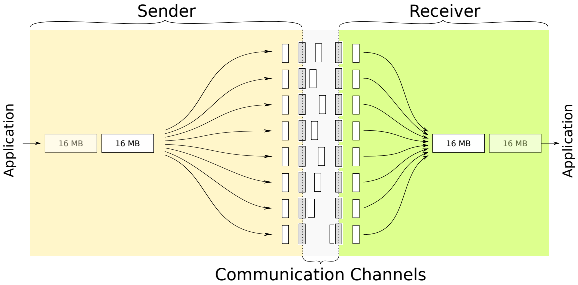
Brandenburg University of Technology Cottbus - Senftenberg



ADAPTATION

FOR LOW LATENCY

Brandenburg University of Technology Cottbus - Senftenberg

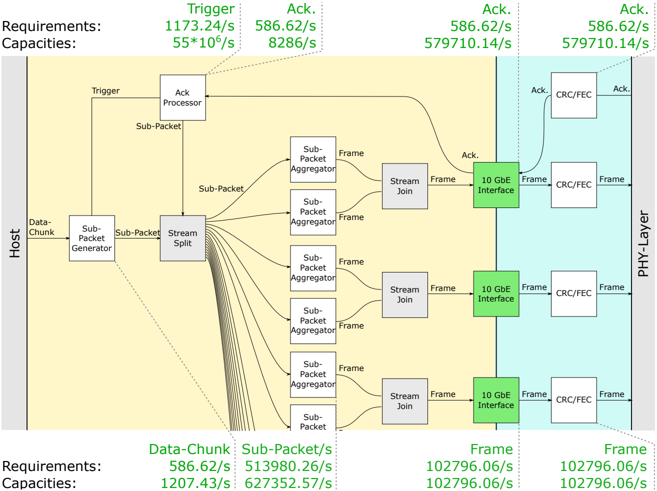


One Data-Chunk is transmitted in parallel:

- Throughput depends on low packet loss
- Protocol management overhead is not hidden
- Latency is expected to decrease accordingly to the number of interfaces

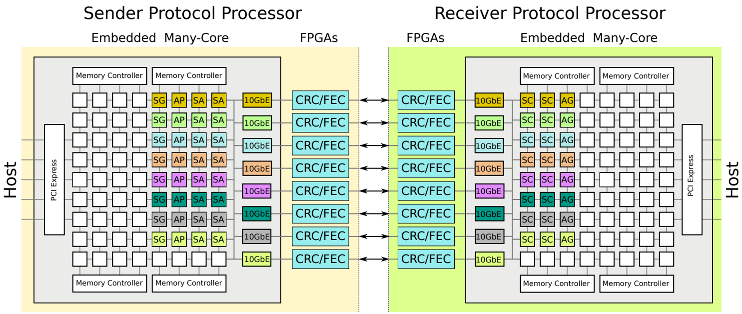
ADAPTATION

PROTOCOL PROCESSING ENGINE FOR LOW LATENCY



MAPPING

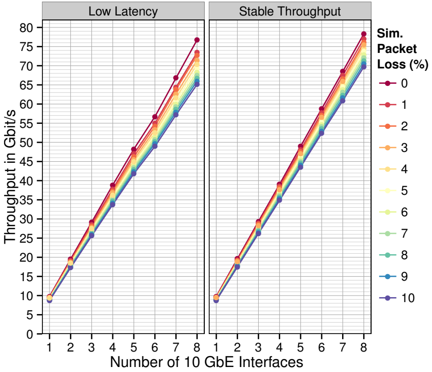
STABLE THROUGHPUT



Evaluation System

- 2x TILEncore Gx72 Manycore – high level protocol processing
- Xilinx Virtex-7 FPGA – offloading of compute intensive protocol parts
- Connected by 10 GbE Interfaces

EVALUATION - THROUGHPUT
NO FEC



Stable Throughput

- 78.315 Gbit/s out of theoretical max. 80 Gbit/s
- Processing/Management Overhead:
  min. 0.495% - max. 0.927%
  0.050 Gbit/s (0.50%) 1x 10GbE and
  0.444 Gbit/s (0.55%) 8x 10GbE

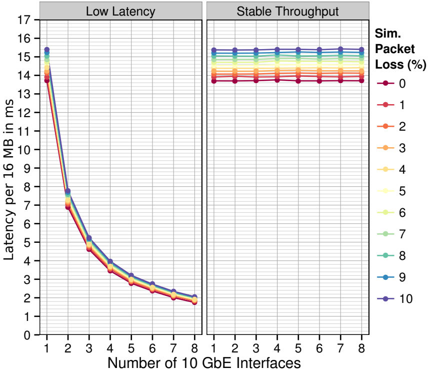
Low Latency

- 76.73 Gbit/s out of theoretical max. 80 Gbit/s
- Processing/Management Overhead:
  min. 0.56% - max. 04.02%
  0.056 Gbit/s (0.56%) 1x 10GbE and
  2.054 Gbit/s (2.57%) 8x 10GbE

Offloaded FEC

- 2.45% lower throughput due to offloading

EVALUATION - LATENCY PER 16MB DATA-CHUNK NO FEC



Stable Throughput

- Almost constant latency

Low Latency

- Latency decreases according to the number of interfaces
- Latency:
  13.714 ms (1x GbE)
  1.748 ms (8x GbE)

⇒ ~34 μs off the optimal scaling

Offloaded FEC

- 2.52% higher latency due to offloading

CONCLUSION

Brandenburg University of Technology Cottbus - Senftenberg

Modeling communication protocols as a soft real-time stream processing problem helps us to . . .

- . . . design highly scalable protocols,
- . . . analyze a protocol's soft real-time requirements,
- . . . adapt a protocol to fit the hardware's capacities,
- . . . and implement the protocol without paradigm changes.

Thank you for your attention

The project End2End100 is a joint project of IHP/BTU and is part to the DFG (German Research Foundation) priority program ”100 Gbit/s Wireless And Beyond”. [2]

[2] DFG Schwerpunktprogram SPP 1655 Drahtlose Ultrahochgeschwindigkeitskommunikation für den mobilen Internetzugriff